

plc

## PATENT ABSTRACTS OF JAPAN

(11)Publication number : 08-116514

(43)Date of publication of application : 07.05.1996

---

(51)Int.Cl.

H04N 5/93

H04N 5/783

---

(21)Application number : 06-252042 (71)Applicant : NIPPON TELEGR & TELEPH  
CORP <NTT>

(22)Date of filing : 18.10.1994 (72)Inventor : MINAMI KENICHI  
AKUTSU AKITO  
HAMADA HIROSHI

---

(54) VIDEO IMAGE SUMMARIZING METHOD AND DEVICE THEREFOR

(57)Abstract:

PURPOSE: To reduce a reproduction time by avoiding unnaturality in sound to be produced when a fast viewing is realized without changing a reproduction speed of a video image.

CONSTITUTION: A video image received by a video image input section 101 in real time is fed to a video image reproduction control section 103 while being stored in a video storage section 102. When the user commands a reproduction method (1) from an interface section 109a cut-point detection section 104 detects a cut-point of the video image. Then a low power detection section 105 detects points at which power of audio information is reduced around the cut points and points at which the power of the audio information gets lower after and before the former points. Then a reproduction time is adjusted and buffering for a video image block is made a video output section 108 reproduces the video image from points at which the power of the audio information gets lower around each cut-point till the points of succeeding audio information and the reproduced image is outputted to a display device of the interface section 109.

---

### CLAIMS

---

[Claim(s)]

[Claim 1]As opposed to said video information photoed in accumulated video information or real time which consists of picture information and speech informationAn image summarizing method which reproduces an image between a point that asked for two or more cut points when a scene of a picture changes from said picture informationand power of speech information became low in the

circumference of these each cut point and a point that power of speech information became low before and behind that one by one.

[Claim 2] As opposed to said video information photoed in accumulated video information or real time which consists of picture information and speech information, an image summarizing method which reproduces an image from a point that asked for two or more cut points when a scene of a picture changes from said picture information and power of speech information became low in the circumference of these each cut point to a certain time one by one.

[Claim 3] As opposed to said video information photoed in accumulated video information or real time which consists of picture information and speech information, an image summarizing method which reproduces an image from a point which only a certain time traced back from a point that asked for two or more cut points when a scene of a picture changes from said picture information and power of speech information became low in the circumference of these each cut point to the point concerned that power of speech information became low one by one.

[Claim 4] The image summarizing method according to claim 3 which raises power as power of speech information corresponding to a point which starts reproduction is made low and it reproduces when reproducing an image.

[Claim 5] The image summarizing method according to claim 2 which makes power of speech information low as a point which ends reproduction is approached when reproducing an image.

[Claim 6] An image abstract device comprising:

An image accumulation part which accumulates a video input section which was photoed in real time and which inputs video information which consists of picture information and speech information and/or said video information.

A cut check appearance part which asks for two or more cut points when a scene of a picture changes from said picture information.

A point that power of speech information became low in the circumference of each of said cut point and an image reproduction control section in which power of speech information includes a low power primary detecting element which detects \*\*\*\*\* low before and behind that.

A point that power of speech information became low in the circumference of each of said cut point, a video output part which power of speech information is low and reproduces an image between \*\*\*\*\*s one by one before and behind that and an interface part which controls setting out of a reproducing starting point of an image, a reproduction end point or regeneration time.

[Claim 7] An image abstract device comprising:

An image accumulation part which accumulates a video input section which was photoed in real time and which inputs video information which consists of picture information and speech information and/or said video information.

A cut check appearance part which asks for two or more cut points when a scene of a picture changes from said picture information.

An image reproduction control section including a low power primary detecting

element which searches for a point that power of speech information became low in the circumference of each of said cut point.

A video output part which reproduces an image from a point that power of speech information became low in the circumference of each of said cut point to a certain time one by one and an interface part which controls setting out of a reproducing starting point of an image reproduction end point or regeneration time.

[Claim 8] An image abstract device comprising:

An image accumulation part which accumulates a video input section which was photoed in real time and which inputs video information which consists of picture information and speech information and/or said video information.

A cut check appearance part which asks for two or more cut points when a scene of a picture changes from said picture information.

An image reproduction control section including a low power primary detecting element which searches for a point that power of speech information became low in the circumference of each of said cut point.

A video output part which reproduces an image from a time of only a certain time going back from a point that power of speech information became low in the circumference of each of said cut point to the point concerned that power of speech information became low one by one and an interface part which controls setting out of a reproducing starting point of an image reproduction end point or regeneration time.

[Claim 9] An image abstract device which has an interface part which controls an image reproduction control section characterized by comprising the following: a video output part which reproduces an image in accordance with a selected regeneration method; directions of a regeneration method of an image; and setting out of a reproducing starting point of an image reproduction end point or regeneration time.

An image accumulation part which accumulates a video input section which was photoed in real time and which inputs video information which consists of picture information and speech information and/or said video information.

A cut check appearance part which asks for two or more cut points when a scene of a picture changes from said picture information.

A low power primary detecting element which searches for a point that power of speech information became low in the circumference of said cut point.

An image between a point that power of speech information became low in the circumference of each of said cut point and a point that power of speech information became low before and behind that one by one. An image by a certain time from the 1st regeneration method to reproduce and a point that power of speech information became low in the circumference of each of said cut point one by one. From the 2nd regeneration method to reproduce and a point which only a certain time traced back from a point that power of speech information became low in the circumference of each of said cut point. A regeneration method

selecting part which at least two methods are beforehand set up among the 3rd regeneration method that reproduces an image to the point concerned that power of speech information became low one by one and chooses a directed regeneration method.

[Claim 10] The image abstract device according to claim 8 or 9 containing a power adjustment part which raises power as said image reproduction control section makes low power of speech information corresponding to a point which starts reproduction when reproducing an image and is reproduced.

[Claim 11] The image abstract device according to claim 7 or 9 containing a power adjustment part which makes power of speech information low as it approaches a point which ends reproduction when said image reproduction control section reproduces an image.

---

## DETAILED DESCRIPTION

---

[Detailed Description of the Invention]

[0001]

[Industrial Application] This invention relates to an image summarizing method and a device.

[0002]

[Description of the Prior Art] Although the abstract art in consideration of the semantic contents of the image does not exist now a certain physical quantity is extracted from an image and the art of shortening the time at the time of taking out the portion which fulfills conditions with physical quantity and seeing an image exists. For example according to the difference quantity for every frame of a picture the variation of luminosity and the variation of a color and a histogram it reproduces at high speed [when there is little change] The cut point of the method (Japanese Patent Application No. 3-20519 "image-contents compression display processing method") of reproducing at the rate of usual when there is much change or an image is detected There are some (Japanese Patent Application No. 5-15422 "method of creating the Hayami image and device") which realize slanting reading of an image by reproducing the image of immediately after a cut point or a just before for several seconds or fixing the length of the image corresponding to one shot and reproducing the image of fixed time. There is art of displaying the picture immediately after a cut point on a display side by side spatially as a still picture or printing on paper and grasping the contents at a glance.

[0003]

[Problem(s) to be Solved by the Invention] In order to acquire required information efficiently out of the overflowing video information it is necessary to see an image for a short time. Although there were some to which the speed reproduced according to the characteristic quantity extracted from picture information is changed conventionally when reproducing at high speed speech information could

not be added but the amount of information for grasping the contents was falling. It became a piece which a sound strikes with the method of reproducing only cut point order in the reproductive starting point and end point when reproducing an image since speech information is not taken into consideration and was unnatural.

[0004] The purpose of this invention cancels the unnaturalness of the sound made when realizing Hayami of an image without changing the reproduction speed of an image and there is in providing the image summarizing method and device which are considered as an image abstract with shortening of regeneration time.

[0005]

[Means for Solving the Problem] As opposed to said video information photoed in accumulated video information or real time when an image summarizing method of claim 1 becomes from picture information and speech information It asks for two or more cut points when a scene of a picture changes from said picture information and an image between a point that power of speech information became low in the circumference of these each cut point and a point that power of speech information became low before and behind that is reproduced one by one.

[0006] As opposed to said video information photoed in accumulated video information or real time when an image summarizing method of claim 2 becomes from picture information and speech information It asks for two or more cut points when a scene of a picture changes from said picture information and an image from a point that power of speech information became low in the circumference of these each cut point to a certain time is reproduced one by one.

[0007] As opposed to said video information photoed in accumulated video information or real time when an image summarizing method of claim 3 becomes from picture information and speech information It asks for two or more cut points when a scene of a picture changes from said picture information and an image from a point which only a certain time traced back from a time of power of speech information becoming low in the circumference of these each cut point to the point concerned that power of speech information became low is reproduced one by one.

[0008] Power of speech information corresponding to a point which starts reproduction is made low and when reproducing an image power can be raised in an image summarizing method of claim 3 as it reproduces.

[0009] In an image summarizing method of claim 2 when reproducing an image power of speech information can be made small as a point which ends reproduction is approached.

[0010] An image accumulation part which accumulates a video input section into which an image abstract device of claim 6 inputs video information which was photoed in real time and which consists of picture information and speech information and/or said video information A cut check appearance part which asks for two or more cut points when a scene of a picture changes from said picture information A point that power of speech information became low in the circumference of each of said cut point and an image reproduction control section in which power of speech information includes a low power primary

detecting element which detects \*\*\*\*\* low before and behind thatIt has a video output part which reproduces an image between a point that power of speech information became low in the circumference of each of said cut pointand a point that power of speech information became low before and behind thatone by oneand an interface part which controls setting out of a reproducing starting point of an imagea reproduction end pointor regeneration time.

[0011]An image accumulation part which accumulates a video input section into which an image abstract device of claim 7 inputs video information which was photoed in real timeand which consists of picture information and speech informationand/or said video informationA cut check appearance part which asks for two or more cut points when a scene of a picture changes from said picture informationAn image reproduction control section including a low power primary detecting element which detects a point that power of speech information became low in the circumference of each of said cut pointIt has a video output part which reproduces an image from a point that power of speech information became low in the circumference of each of said cut point to a certain timeone by oneand an interface part which controls setting out of a reproducing starting point of an imagea reproduction end pointor regeneration time.

[0012]An image accumulation part which accumulates a video input section into which an image abstract device of claim 8 inputs video information which was photoed in real timeand which consists of picture information and speech informationand/or said video informationA cut check appearance part which asks for two or more cut points when a scene of a picture changes from said picture informationAn image reproduction control section including a low power primary detecting element which searches for a point that power of speech information became low in the circumference of each of said cut pointA video output part which reproduces an image from a time of only a certain time going back from a point that power of speech information became low in the circumference of each of said cut point to the point concerned that power of speech information became lowone by oneIt has an interface part which controls setting out of a reproducing starting point of an imagea reproduction end pointor regeneration time.

[0013]An image accumulation part which accumulates a video input section into which an image abstract device of claim 9 inputs video information which was photoed in real timeand which consists of picture information and speech informationand/or said video informationA cut check appearance part which asks for two or more cut points when a scene of a picture changes from said picture informationA low power primary detecting element which detects a point that power of speech information became low in the circumference of each of said cut pointAn image between a point that power of speech information became low in the circumference of each of said cut pointand a point that power of speech information became low before and behind that one by oneAn image by a certain time from the 1st regeneration method to reproduce and a point that power of speech information became low in the circumference of each of said cut point one by oneFrom the 2nd regeneration method to reproduce and a point which only a

certain time traced back from a point that power of speech information became low in the circumference of each of said cut point. An image reproduction control section in which at least two methods are beforehand set and contain a regeneration method selecting part which chooses a directed regeneration method among the 3rd regeneration method that reproduces an image to the point concerned that power of speech information became low by one A video output part which reproduces an image in accordance with a selected regeneration method directions of a regeneration method of an image and a reproducing starting point of an image. It has an interface part which controls setting out of a reproduction end point or regeneration time.

[0014] In an image abstract device of claim 8 or 9 when an image reproduction control section reproduces an image it can make low power of speech information corresponding to a point which starts reproduction and can contain a power adjustment part to reproduce and which is a like therefore raises power.

[0015] In an image abstract device of claim 7 or 9 when an image reproduction control section reproduces an image it can contain a power adjustment part which makes power of speech information low as it approaches a point which ends reproduction.

[0016]

[Function] The image summarizing method of claim 1 can cancel the piece which an unnatural sound beats when the power of speech information uses for the reproducing starting point and end point of an image the portion which became low.

[0017] The image summarizing method of claim 2 can cancel the piece which the unnatural sound in the reproducing starting point of an image beats when the power of speech information uses for the reproducing starting point of an image the portion which became low.

[0018] The image summarizing method of claim 3 can cancel the piece which the unnatural sound in the reproduction end point of an image beats when the power of speech information uses for the reproduction end point of an image the portion which became low.

[0019] Since the length of each reproducing section can be determined arbitrarily the image summarizing method of claims 2 and 3 is effective when regeneration time is restricted.

[0020] When reproducing an image the image summarizing method of claim 4 makes low power of the speech information corresponding to the point which starts reproduction and it can cancel the piece which the unnatural sound in the reproducing starting point of an image beats by raising power as it reproduces.

[0021] When the image summarizing method of claim 5 reproduces an image it can cancel the piece which the unnatural sound in the reproduction end point of an image beats by making power of speech information low as it approaches the point which ends reproduction.

[0022] The image abstract device of claims 6 and 7 is a device which enforces the image summarizing method of claims 2 and 3 respectively.

[0023] The abstract device of claim 9 enables it to choose one regeneration

method from two or more regeneration methods with one device.

[0024]The image abstract device of claims 10 and 11 is a device which enforces the image summarizing method of claims 4 and 5 respectively.

[0025]

[Example]Next the example of this invention is described with reference to drawings.

[0026]The flow chart with which drawing 1 shows the block diagram of the image abstract device of one example of this invention and drawing 2 shows the flow of processing of the image reproduction control section 103 and drawing 3 are the figures showing an example of the interface part 109.

[0027]The video input section 101 into which the image abstract device of this example inputs an image and the image accumulation part 102 which accumulates the image inputted in real time and outputs the accumulated image in real time. It comprises the image reproduction control section 103 which controls the image to reproduce the video output part 108 which reproduces an image and is outputted and the interface part 109 to which a user points to a regeneration method and a reproducing starting point, a reproduction end point and regeneration time and shows a user the reproduced image. The cut check appearance part 104 from which the image reproduction control section 103 detects the cut point when the scene (shot) of an image changes. The power of the speech information of an image comprises the power adjustment part 106 which adjusts the power of the speech information of the low power primary detecting element 105 which detects a low portion, the point which starts reproduction and the point to end and the regeneration method selecting part 107 which chooses regeneration time and a regeneration method. The interface part 109 comprises the display 401 which projects an image, regeneration method selection button 402<sub>1</sub> - 402<sub>3</sub> and the regeneration time selection knob 403.

[0028]Here the following three kinds are set as the regeneration method selecting part 107 as a regeneration method.

- (1) Reproduce the image between the point that the power of speech information became low in the circumference of each cut point of an image and the point that the power of speech information became low before and behind that one by one.
- (2) Reproduce a certain time and an image one by one from the time of the power of speech information becoming low in the circumference of each cut point of an image.
- (3) Reproduce the image from the point which only a certain time traced back from the time of the power of speech information becoming low in the circumference of each cut point of an image to the point concerned that the power of speech information became low one by one.

[0029]Next operation of this example is explained.

[0030]The image inputted by the video input section 101 in real time is sent to the image reproduction control section 103 being accumulated in the image accumulation part 102. In the case of the image accumulated beforehand it is directly sent from the image accumulation part 102 at the image reproduction



control section 103. Next processing shown by drawing 2 is performed based on the regeneration method to which the user pointed by the interface part 109. First cut check appearance processing 201 is performed to an image by the cut check appearance part 104. The method by "the video-cuts check attitude method and device" of Japanese Patent Application No. 5-317663 etc. are used for detection of a cut point. Next low power detection processing 202 is performed in the low power primary detecting element 105 and the point that the power of speech information became low in the circumference of a cut point is searched for. The value of the sum of squares of the voice waveform in every several ms etc. are used for detection of low power. Since regeneration time cannot specify strictly when using the low power part of speech information processing is performed so that a low power part which becomes the closest to the specified regeneration time may be used. In not using a low power part for the end of reproduction of an image and a reproducing starting point (a regeneration method (2)(3)) weighting as shown in 301 and 302 of drawing 3 is performed in the power adjustment part 106 (power adjustment processing 203) and it reduces the piece which an unnatural sound beats. Although power adjustment is performed with the primary function at drawing 3 it is also possible to attach a smooth change like a sigmoid function. After reproduction pretreatment 204 of adjustment of regeneration time buffering of each image section etc. is performed by the regeneration method selecting part 107 an image is reproduced in the video output part 108. The reproduced image is projected on the display 401 of the interface part 109.

[0031] Although it enabled it to choose two or more regeneration methods in this example it may be made to perform only any one regeneration method. It may be made to include the display 401 in the video output part 108. At least either of the video input section 101 and the image accumulation part 102 is good.

[0032]

[Effect of the Invention] As explained above this invention has an effect as taken below.

- (1) The invention of claims 16 and 9 can cancel the piece which the unnatural sound in the reproducing starting point and end point of an image beats when the power of speech information uses for the reproducing starting point and end point of an image the portion which became low.
- (2) The invention of claims 27 and 9 can cancel the piece which the unnatural sound in the reproducing starting point of an image beats when the power of speech information uses for the reproducing starting point of an image the portion which became low.
- (3) The invention of claims 38 and 9 can cancel the piece which the unnatural sound in the reproduction end point of an image beats when the power of the speech information of the reproduction end point of an image uses the portion which became low.
- (4) When the invention of claims 4 and 10 reproduces an image it can cancel the piece which the unnatural sound in the reproducing starting point of an image

beats by raising power as power of the speech information corresponding to the point which starts reproduction is made low and it reproduces.

(5) When the invention of claims 5 and 11 reproduces an image it can cancel the piece which the unnatural sound in the reproduction end point of an image beats by making power of speech information low as it approaches the point which ends reproduction.

---

## DESCRIPTION OF DRAWINGS

---

[Brief Description of the Drawings]

[Drawing 1] It is a block diagram showing the outline composition of the image abstract device of one example of this invention.

[Drawing 2] It is a flow chart which shows the flow of processing of the image reproduction control section 104 of drawing 1.

[Drawing 3] It is a figure showing a situation when weighting is given to the power of the speech information of the reproducing starting point by this example and an end point.

[Drawing 4] It is a figure showing the composition of the interface part 109 of drawing 1.

[Description of Notations]

101 Video input section

102 Image accumulation part

103 Image reproduction control section

104 Cut check appearance part

105 Low power primary detecting element

106 Power adjustment part

107 Regeneration method selecting part

108 Video output part

109 Interface part

201 Cut check appearance processing

202 Low power detection processing

203 Power adjustment processing

204 Reproduction pretreatment

301 302 weighting

401 Display

402<sub>1</sub> - a 402<sub>3</sub> regeneration method selection button

403 Regeneration time selection knob

---

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平8-116514

(43) 公開日 平成8年(1996)5月7日

(51) Int.Cl. <sup>s</sup>	識別記号	庁内整理番号	F I	技術表示箇所
H 0 4 N 5/93				
5/783	G		H 0 4 N 5/ 93	Z
				G

審査請求 未請求 請求項の数11 O L (全 6 頁)

(21) 出願番号 特願平6-252042

(22) 出願日 平成6年(1994)10月18日

(71) 出願人 000004226

日本電信電話株式会社  
東京都新宿区西新宿三丁目19番2号

(72) 発明者 南 憲一

東京都千代田区内幸町一丁目1番6号 日  
本電信電話株式会社内

(72) 発明者 阿久津 明人

東京都千代田区内幸町一丁目1番6号 日  
本電信電話株式会社内

(72) 発明者 浜田 洋

東京都千代田区内幸町一丁目1番6号 日  
本電信電話株式会社内

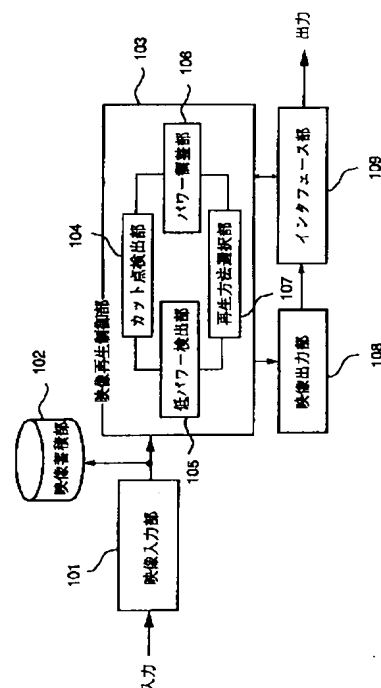
(74) 代理人 弁理士 若林 忠

(54) 【発明の名称】 映像要約方法および装置

(57) 【要約】

【目的】 映像の再生速度を変化させることなく映像の速見を実現する場合に生じる音の不自然さを解消し、再生時間を短縮する。

【構成】 実時間で映像入力部101によって入力された映像は、映像蓄積部102に蓄積されながら映像再生制御部103に送られる。インタフェース部109からユーザーが再生方法(1)を指示すると、カット点検出部105によって映像のカット点が検出される。次に、低パワー検出部105によって、これら各カット点の周辺において音声情報のパワーが低くなった点とその前後で音声情報のパワーが低くなった点検出される。次に、再生時間の調整や映像区間のバッファリングが行われた後、映像出力部108によって、各カット点の周辺において音声情報のパワーが低くなった点から低くなった点までの映像が再生され、インタフェース部109のディスプレイに出力される。



**【特許請求の範囲】**

【請求項 1】 画像情報と音声情報からなる蓄積された映像情報あるいは実時間で撮影された前記映像情報に対して、前記画像情報から画像の場面が変化する複数のカット点を求め、これら各カット点の周辺において音声情報のパワーが低くなった点とその前後で音声情報のパワーが低くなった点の間の映像を順次、再生する映像要約方法。

【請求項 2】 画像情報と音声情報からなる蓄積された映像情報あるいは実時間で撮影された前記映像情報に対して、前記画像情報から画像の場面が変化する複数のカット点を求め、これら各カット点の周辺において音声情報のパワーが低くなった点からある時間までの映像を順次、再生する映像要約方法。

【請求項 3】 画像情報と音声情報からなる蓄積された映像情報あるいは実時間で撮影された前記映像情報に対して、前記画像情報から画像の場面が変化する複数のカット点を求め、これら各カット点の周辺において音声情報のパワーが低くなった点からある時間だけさかのぼった点から、音声情報のパワーが低くなった当該点までの映像を順次、再生する映像要約方法。

【請求項 4】 映像を再生する際に、再生を開始する点に対応する音声情報のパワーを低くし、再生するにしたがってパワーを上げる請求項 3 に記載の映像要約方法。

【請求項 5】 映像を再生する際に、再生を終了する点に近づくにしたがって音声情報のパワーを低くする請求項 2 に記載の映像要約方法。

【請求項 6】 実時間で撮影された、画像情報と音声情報とからなる映像情報を入力する映像入力部および／または前記映像情報を蓄積する映像蓄積部と、前記画像情報から画像の場面が変化するカット点を複数求めるカット点検出部と、前記各カット点の周辺において音声情報のパワーが低くなった点とその前後で音声情報のパワーが低くなった点を検出する低パワー検出部を含む映像再生制御部と、前記各カット点の周辺において音声情報のパワーが低くなった点とその前後で音声情報のパワーが低くなった点の間の映像を順次、再生する映像出力部と、映像の再生開始点、再生終了点や再生時間の設定を制御するインタフェース部を有する映像要約装置。

【請求項 7】 実時間で撮影された、画像情報と音声情報とからなる映像情報を入力する映像入力部および／または前記映像情報を蓄積する映像蓄積部と、前記画像情報から画像の場面が変化するカット点を複数求めるカット点検出部と、前記各カット点の周辺において音声情報のパワーが低くなった点を求める低パワー検出部を含む映像再生制御部と、前記各カット点の周辺において音声情報のパワーが低くなった点からある時間までの映像を順次、再生する映像出力部と、

映像の再生開始点、再生終了点や再生時間の設定を制御するインタフェース部を有する映像要約装置。

【請求項 8】 実時間で撮影された、画像情報と音声情報とからなる映像情報を入力する映像入力部および／または前記映像情報を蓄積する映像蓄積部と、前記画像情報から画像の場面が変化するカット点を複数求めるカット点検出部と、前記各カット点の周辺において音声情報のパワーが低くなった点を求める低パワー検出部を含む映像再生制御部と、前記各カット点の周辺において音声情報のパワーが低くなった点からある時間だけさかのぼった時点から、音声情報のパワーが低くなった当該点までの映像を順次、再生する映像出力部と、

映像の再生開始点、再生終了点や再生時間の設定を制御するインタフェース部を有する映像要約装置。

【請求項 9】 実時間で撮影された、画像情報と音声情報とからなる映像情報を入力する映像入力部および／または前記映像情報を蓄積する映像蓄積部と、前記画像情報から画像の場面が変化するカット点を複数求めるカット点検出部と、前記カット点の周辺において音声情報のパワーが低くなった点を求める低パワー検出部と、前記各カット点の周辺において音声情報のパワーが低くなった点とその前後で音声情報のパワーが低くなった点の間の映像を順次、再生する第 1 の再生方法、前記各カット点の周辺において音声情報のパワーが低くなった点からある時間までの映像を順次、再生する第 2 の再生方法、前記各カット点の周辺において音声情報のパワーが低くなった点からある時間だけさかのぼった点から、音声情報のパワーが低くなった当該点までの映像を順次、再生する第 3 の再生方法のうち少なくとも 2 つの方法が予め設定され、指示された再生方法を選択する再生方法選択部とを含む映像再生制御部と、選択された再生方法にしたがって映像を再生する映像出力部と、

映像の再生方法の指示と映像の再生開始点、再生終了点や再生時間の設定を制御するインタフェース部を有する映像要約装置。

【請求項 10】 前記映像再生制御部は、映像を再生する際に、再生を開始する点に対応する音声情報のパワーを低くし、再生するにしたがってパワーを上げるパワー調整部を含む請求項 8 または 9 に記載の映像要約装置。

【請求項 11】 前記映像再生制御部は、映像を再生する際に、再生を終了する点に近づくにしたがって音声情報のパワーを低くするパワー調整部を含む請求項 7 または 9 に記載の映像要約装置。

【発明の詳細な説明】

【0001】

【産業上の利用分野】 本発明は映像要約方法および装置に関する。

【0002】

【従来の技術】映像の意味的な内容を考慮した要約技術は現在存在しないが、映像からある物理量を抽出し、物理量がある条件を満たす部分を取り出して映像を見る際の時間を短縮する技術は存在する。例えば、画像のフレーム毎の差分量や輝度の変化量、色とヒストグラムの変化量に応じて変化の少ない場合には高速に再生し、変化の多い場合には通常で再生する方法（特願平3-20519「映像内容圧縮表示処理方法」）や映像のカット点を検出し、カット点の直後や直前の映像を数秒間再生したり、1ショットに対応する映像の長さを固定して一定時間の映像を再生することで映像の斜め読みを実現するもの（特願平5-15422「速見映像を作成する方法および装置」）がある。また、カット点直後の画像を静止画として空間的に並べてディスプレイに表示したり、紙に印刷して内容を一目で把握する技術がある。

【0003】

【発明が解決しようとする課題】氾濫する映像情報の中から必要な情報を効率よく得るためには、映像を短時間で見る必要がある。従来、画像情報から抽出した特徴量に応じて再生する速度を変化させるものがあったが、高速に再生する際には音声情報を付加することができず、内容を把握するための情報量が低下していた。また、カット点の前後のみを再生する方法では、音声情報を考慮していないため、映像を再生する際に再生の開始点と終了点において音がぶつ切れになり、不自然であった。

【0004】本発明の目的は、映像の再生速度を変化させることなく映像の速見を実現する場合に生じる音の不自然さを解消し、再生時間の短縮をもって映像要約とする映像要約方法および装置を提供することにある。

【0005】

【課題を解決するための手段】請求項1の映像要約方法は、画像情報と音声情報からなる蓄積された映像情報あるいは実時間で撮影された前記映像情報に対して、前記画像情報から画像の場面が変化する複数のカット点を求め、これら各カット点の周辺において音声情報のパワーが低くなった点とその前後で音声情報のパワーが低くなった点の間の映像を順次、再生する。

【0006】請求項2の映像要約方法は、画像情報と音声情報からなる蓄積された映像情報あるいは実時間で撮影された前記映像情報に対して、前記画像情報から画像の場面が変化する複数のカット点を求め、これら各カット点の周辺において音声情報のパワーが低くなった点からある時間までの映像を順次、再生する。

【0007】請求項3の映像要約方法は、画像情報と音声情報からなる蓄積された映像情報あるいは実時間で撮影された前記映像情報に対して、前記画像情報から画像の場面が変化する複数のカット点を求め、これら各カット点の周辺において音声情報のパワーが低くなった時点からある時間だけさかのぼった点から、音声情報のパワーが低くなった当該点までの映像を順次、再生する。

【0008】請求項3の映像要約方法では、映像を再生する際に、再生を開始する点に対応する音声情報のパワーを低くし、再生するにしたがってパワーを上げることができる。

【0009】請求項2の映像要約方法では、映像を再生する際、再生を終了する点に近づくにしたがって音声情報のパワーを小さくすることができる。

【0010】請求項6の映像要約装置は、実時間で撮影された、画像情報と音声情報とからなる映像情報を入力する映像入力部および／または前記映像情報を蓄積する映像蓄積部と、前記画像情報から画像の場面が変化するカット点を複数求めるカット点検出部と、前記各カット点の周辺において音声情報のパワーが低くなった点とその前後で音声情報のパワーが低くなった点を検出する低パワー検出部を含む映像再生制御部と、前記各カット点の周辺において音声情報のパワーが低くなった点とその前後で音声情報のパワーが低くなった点の間の映像を順次、再生する映像出力部と、映像の再生開始点、再生終了点や再生時間の設定を制御するインタフェース部を有する。

【0011】請求項7の映像要約装置は、実時間で撮影された、画像情報と音声情報とからなる映像情報を入力する映像入力部および／または前記映像情報を蓄積する映像蓄積部と、前記画像情報から画像の場面が変化するカット点を複数求めるカット点検出部と、前記各カット点の周辺において音声情報のパワーが低くなった点を検出する低パワー検出部を含む映像再生制御部と、前記各カット点の周辺において音声情報のパワーが低くなった点からある時間までの映像を順次、再生する映像出力部と、映像の再生開始点、再生終了点や再生時間の設定を制御するインタフェース部を有する。

【0012】請求項8の映像要約装置は、実時間で撮影された、画像情報と音声情報とからなる映像情報を入力する映像入力部および／または前記映像情報を蓄積する映像蓄積部と、前記画像情報から画像の場面が変化するカット点を複数求めるカット点検出部と、前記各カット点の周辺において音声情報のパワーが低くなった点を求める低パワー検出部を含む映像再生制御部と、前記各カット点の周辺において音声情報のパワーが低くなった点からある時間だけさかのぼった時点から、音声情報のパワーが低くなった当該点までの映像を順次、再生する映像出力部と、映像の再生開始点、再生終了点や再生時間の設定を制御するインタフェース部を有する。

【0013】請求項9の映像要約装置は、実時間で撮影された、画像情報と音声情報とからなる映像情報を入力する映像入力部および／または前記映像情報を蓄積する映像蓄積部と、前記画像情報から画像の場面が変化するカット点を複数求めるカット点検出部と、前記各カット点の周辺において音声情報のパワーが低くなった点を検出する低パワー検出部と、前記各カット点の周辺におい

て音声情報のパワーが低くなった点とその前後で音声情報のパワーが低くなった点の間の映像を順次、再生する第 1 の再生方法、前記各カット点の周辺において音声情報のパワーが低くなった点からある時間までの映像を順次、再生する第 2 の再生方法、前記各カット点の周辺において音声情報のパワーが低くなった点からある時間だけさかのぼった点から、音声情報のパワーが低くなった当該点までの映像を順次、再生する第 3 の再生方法のうち少なくとも 2 つの方法が予め設定され、指示された再生方法を選択する再生方法選択部とを含む映像再生制御部と、選択された再生方法にしたがって映像を再生する映像出力部と、映像の再生方法の指示と映像の再生開始点、再生終了点や再生時間の設定を制御するインタフェース部を有する。

【0014】請求項 8 または 9 の映像要約装置では、映像再生制御部は、映像を再生する際に、再生を開始する点に対応する音声情報のパワーを低くし、再生するにしたがってパワーを上げるパワー調整部を含むことができる。

【0015】請求項 7 または 9 の映像要約装置では、映像再生制御部は、映像を再生する際に、再生を終了する点に近づくにしたがって音声情報のパワーを低くするパワー調整部を含むことができる。

【0016】

【作用】請求項 1 の映像要約方法は、映像の再生開始点と終了点に音声情報のパワーが低くなった部分を用いることにより、不自然な音のぶつ切れが解消できる。

【0017】請求項 2 の映像要約方法は、映像の再生開始点に音声情報のパワーが低くなった部分を用いることにより、映像の再生開始点における不自然な音のぶつ切れを解消できる。

【0018】請求項 3 の映像要約方法は、映像の再生終了点に音声情報のパワーが低くなった部分を用いることにより、映像の再生終了点における不自然な音のぶつ切れを解消できる。

【0019】請求項 2 および 3 の映像要約方法は、個々の再生区間の長さを任意に決定できるため、再生時間が限られている場合に有効である。

【0020】請求項 4 の映像要約方法は、映像を再生する際に、再生を開始する点に対応する音声情報のパワーを低くし、再生するにしたがってパワーを上げることにより映像の再生開始点における不自然な音のぶつ切れを解消できる。

【0021】請求項 5 の映像要約方法は、映像を再生する際に、再生を終了する点に近づくにしたがって音声情報のパワーを低くすることにより、映像の再生終了点における不自然な音のぶつ切れを解消できる。

【0022】請求項 6, 7, 8 の映像要約装置は、それぞれ請求項 1, 2, 3 の映像要約方法を実施する装置である。

【0023】請求項 9 の要約装置は、一台の装置で複数の再生方法から 1 つの再生方法を選択できるようにしたものである。

【0024】請求項 10, 11 の映像要約装置はそれぞれ請求項 4, 5 の映像要約方法を実施する装置である。

【0025】

【実施例】次に、本発明の実施例について図面を参照して説明する。

【0026】図 1 は本発明の一実施例の映像要約装置のブロック図、図 2 は映像再生制御部 103 の処理の流れを示すフローチャート、図 3 はインタフェース部 109 の一例を示す図である。

【0027】本実施例の映像要約装置は、映像を入力する映像入力部 101 と、実時間で入力された映像を蓄積し、蓄積された映像を実時間で出力する映像蓄積部 102 と、再生する映像を制御する映像再生制御部 103 と、映像を再生し、出力する映像出力部 108 と、ユーザーが再生方法と、再生開始点、再生終了点や再生時間を指示し、また再生された映像をユーザーに提示するインタフェース部 109 で構成されている。映像再生制御部 103 は、映像の場面（ショット）の変化するカット点を検出するカット点検出部 104 と、映像の音声情報のパワーが低い部分を検出する低パワー検出部 105 と、再生を開始する点と終了する点の音声情報のパワーを調整するパワー調整部 106 と、再生時間や再生方法を選択する再生方法選択部 107 で構成されている。インタフェース部 109 は、映像を映し出すディスプレイ 401 と、再生方法選択ボタン 402<sub>1</sub>~402<sub>3</sub>と、再生時間選択つまみ 403 から構成されている。

【0028】ここで、再生方法として、次の 3 通りが再生方法選択部 107 に設定されている。

(1) 映像の各カット点の周辺において音声情報のパワーが低くなった点とその前後で音声情報のパワーが低くなった点の間の映像を順次、再生する。

(2) 映像の各カット点の周辺において音声情報のパワーが低くなった時点からある時間、映像を順次、再生する。

(3) 映像の各カット点の周辺において音声情報のパワーが低くなった時点からある時間だけさかのぼった点から、音声情報のパワーが低くなった当該点までの映像を順次、再生する。

【0029】次に、本実施例の動作を説明する。

【0030】実時間で映像入力部 101 によって入力された映像は、映像蓄積部 102 に蓄積されながら、映像再生制御部 103 に送られる。予め蓄積された映像の場合には映像蓄積部 102 より直接映像再生制御部 103 に送られる。次に、インタフェース部 109 によってユーザーが指示した再生方法に基づき、図 2 で示した処理が施される。まず、カット点検出部 104 によって映像にカット点検出処理 201 が施される。カット点の検出

には、特願平 5-317663 の「映像カット点検出方法および装置」による方法等が用いられる。次に、低パワー検出部 105 において低パワー検出処理 202 が施され、カット点の周辺において音声情報のパワーが低くなった点が求められる。低パワーの検出には、数 ms 毎の音声波形の二乗和の値等が用いられる。音声情報の低パワー部を用いる場合には、再生時間が厳密には指定できないため、指定された再生時間に最も近くなるような低パワー部を用いるように処理が施される。映像の再生終了、再生開始点に低パワー部を用いない場合（再生方法（2）、（3））には、図 3 の 301 および 302 に示したような重み付けをパワー調整部 106 で行い（パワー調整処理 203）、不自然な音のぶつ切れを軽減する。図 3 では 1 次関数でパワー調整を行っているが、シグモイド関数のように滑らかな変化を付けることも可能である。再生方法選択部 107 で再生時間の調整や各映像区間のバッファリング等の再生前処理 204 が行われた後、映像出力部 108 において映像が再生される。再生された映像はインタフェース部 109 のディスプレイ 401 に映し出される。

【0031】なお、本実施例では複数の再生方法を選択できるようにしたが、いずれか 1 つの再生方法だけを行なうようにしてもよい。また、ディスプレイ 401 を映像出力部 108 に含めるようにしてもよい。さらに、映像入力部 101 と映像蓄積部 102 はいずれか一方だけでもよい。

#### 【0032】

【発明の効果】以上説明したように、本発明は以下に示すような効果がある。

（1）請求項 1、6 および 9 の発明は、映像の再生開始点と終了点に音声情報のパワーが低くなった部分を用いることにより、映像の再生開始点と終了点における不自然な音のぶつ切れが解消できる。

（2）請求項 2、7 および 9 の発明は、映像の再生開始点に音声情報のパワーが低くなった部分を用いることにより、映像の再生開始点における不自然な音のぶつ切れを解消できる。

（3）請求項 3、8 および 9 の発明は、映像の再生終了点の音声情報のパワーが低くなった部分を用いることに

より、映像の再生終了点における不自然な音のぶつ切れを解消できる。

（4）請求項 4 および 10 の発明は、映像を再生する際に、再生を開始する点に対応する音声情報のパワーを低くし、再生するにしたがってパワーを上げることにより、映像の再生開始点における不自然な音のぶつ切れを解消できる。

（5）請求項 5 および 11 の発明は、映像を再生する際に、再生を終了する点に近づくにしたがって音声情報のパワーを低くすることにより、映像の再生終了点における不自然な音のぶつ切れを解消できる。

#### 【図面の簡単な説明】

【図 1】本発明の一実施例の映像要約装置の概略構成を示すブロック図である。

【図 2】図 1 の映像再生制御部 104 の処理の流れを示すフローチャートである。

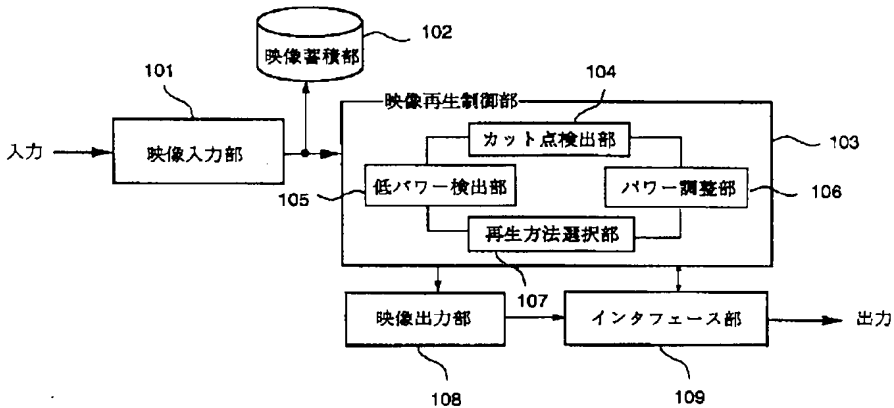
【図 3】本実施例による再生開始点と終了点の音声情報のパワーに重み付けを施したときの様子を示す図である。

【図 4】図 1 のインタフェース部 109 の構成を示す図である。

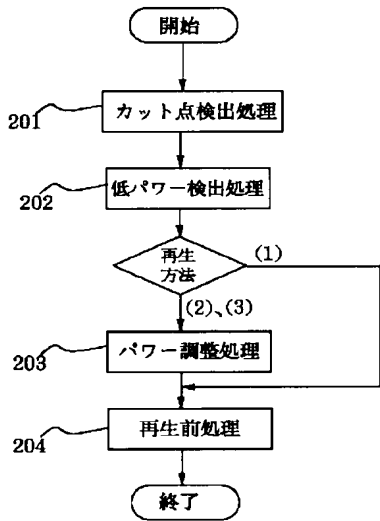
#### 【符号の説明】

101	映像入力部
102	映像蓄積部
103	映像再生制御部
104	カット点検出部
105	低パワー検出部
106	パワー調整部
107	再生方法選択部
108	映像出力部
109	インタフェース部
201	カット点検出処理
202	低パワー検出処理
203	パワー調整処理
204	再生前処理
301, 302	重み付け
401	ディスプレイ
402 <sub>1</sub> ~402 <sub>3</sub>	再生方法選択ボタン
403	再生時間選択つまみ

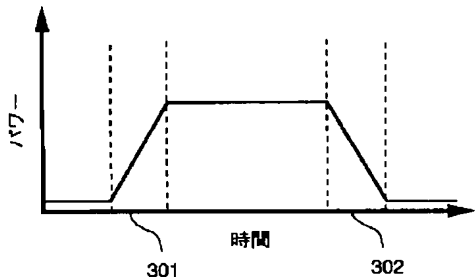
【図 1】



【図 2】



【図 3】



【図 4】

